

DOCUMENTS DE TRAVAIL - WORKING PAPERS

W.P. 07-26

Validité théorique de la Méthode des Choix Discrets :
le cas du Traitement Hormonal substitutif de la Ménopause

Florence Nguyen, Nora Moumjid, Alain Brémond, Marie-Odile Carrère

Novembre 2007

GATE Groupe d'Analyse et de Théorie Économique
UMR 5824 du CNRS
93 chemin des Mouilles – 69130 Écully – France
B.P. 167 – 69131 Écully Cedex
Tél. +33 (0)4 72 86 60 60 – Fax +33 (0)4 72 86 60 90
Messagerie électronique gate@gate.cnrs.fr
Serveur Web : www.gate.cnrs.fr

**Validité théorique de la Méthode des Choix Discrets :
le cas du Traitement Hormonal substitutif de la Ménopause**

Theoretical validity of the Discrete Choice Experiments:
the case of Hormone Replacement Therapy

Florence Nguyen ^{1,2}, Nora Moumjid ², Alain Brémond ², Marie-Odile Carrère ²

²GATE (Groupe d'Analyse et de Théorie Economique), UMR 5824 – CNRS ;
Université Lyon1 ; Université Lyon2 ; ENS LSH ; Centre Léon Bérard, Lyon, France

¹ Florence Nguyen, GRESAC - Centre Léon Bérard, 28 rue Laënnec, 69008 Lyon France,
Email: nguyenf@lyon.fnclcc.fr, **Fax:** (00 33) 4 78 78 28 04, **Tel:** (00 33) 4 78 78 27 80

RESUME

Contexte/Objectif: La Méthode des Choix Discrets (MCD) est utilisée pour analyser les choix des consommateurs. Sous certaines hypothèses de comportement, elle permet d'explicitier les arbitrages des individus entre les attributs d'un bien ou d'un service.

Le Traitement Hormonal substitutif de la Ménopause (THM) est caractérisé par des risques et bénéfices de santé et par un coût monétaire. Nous explorons dans cette situation la capacité de la MCD à révéler les préférences individuelles, en commençant par étudier sa validité théorique. Les comportements sont-ils conformes aux hypothèses standard : cohérence interne, transitivité et stabilité des préférences, absence de préférences lexicographiques, absence d'effet de présentation ? **Méthode:** Onze paires de scénarios ont été construites à partir d'attributs du THM : troubles climatériques, fractures, cancer colorectal, cancer du sein, risque cardiaque, risque thromboembolique, coût. Les variables sociodémographiques et médicales des femmes ont été collectées. Deux versions du questionnaire ont été développées pour tester les effets de présentation des probabilités. La validité théorique a été testée à l'aide de statistiques descriptives. **Résultats:** 462 femmes de la Région Rhône-Alpes âgées de 45 à 65 ans ont renvoyé leur questionnaire. La cohérence interne, la transitivité, et la stabilité des préférences sont observées dans de larges proportions (95%, 97% et 88% des femmes respectivement.). Seules 12% des femmes expriment des préférences lexicographiques. Aucun effet de présentation sur la qualité des réponses n'est détecté. **Conclusion:** Compte tenu de la bonne conformité des comportements aux hypothèses, il sera possible d'estimer une fonction d'utilité et de discuter les préférences des femmes pour le THM ainsi que leur disposition à payer.

Mots-Clés : Méthode des choix discrets ; Révélation des préférences individuelles en santé ; Hypothèses de comportement ; Traitement Hormonal Substitutif de la ménopause.

SUMMARY

Background/Purpose: Discrete Choice Experiments (DCE) consists in providing individuals with fictive scenarios in which characteristics of the good to be evaluated are varying, and asking them to indicate their preferences. Under some behavioural hypotheses, DCE allows to explain individuals' trade-offs between those characteristics. Since Hormone Replacement Therapy (HRT) presents risks, benefits and monetary costs we aim at exploring the potentialities of DCE regarding individual preferences elicitation, by an application to HRT. In this paper, we study the theoretical validity of the method by testing standard behavioural hypotheses: internal consistency, transitivity, and stability of the preferences, no lexicographic preferences, no framing effects. **Methods:** Eleven pairs of scenarios were derived from the following HRT attributes: climacteric troubles, osteoporosis fractures, colorectal cancer, breast cancer, cardiac risk, thromboembolism risk, and monetary cost of the treatment. Women's demographics and medical background were also collected. Two versions of the questionnaire were developed and randomly administered allowing to test framing effects due to the presentation of probabilities. Theoretical validity was tested by descriptive statistics. **Results:** 462 women aged 45 to 65 and from Lyon, France, sent back their questionnaires. Internal consistency, transitivity and stability of preferences were observed in large proportions of women (95%, 97% and 88%, respectively). Only 12% expressed lexicographic preferences. No framing effect was detected on data quality. **Conclusion:** since behavioural hypotheses were verified, it will now be possible to estimate a utility function and to discuss women's preferences and willingness to pay for HRT.

Keywords Discrete Choice Experiments; Individual preferences elicitation; Behavioural hypotheses; Hormone replacement Therapy.

Classification JEL I19 ; D61 ; D83

INTRODUCTION

Dans les économies occidentales, l'allocation des ressources collectives dans le secteur des soins de santé représente un enjeu majeur lié à la survie des régimes de sécurité sociale. Les économistes qui travaillent à cette question essentielle ont proposé différentes méthodes et différents critères de jugement pour la prise en charge ou non des activités de santé par la collectivité. Parmi les approches *welfaristes*, qui consistent à estimer la disposition à payer (DAP) collective pour une activité de santé, deux méthodes sont expérimentées aujourd'hui : la méthode de l'évaluation contingente, depuis le début des années 90, et l'analyse conjointe, plus récemment et surtout dans les pays anglo-saxons. Aujourd'hui, les deux méthodes sont considérées comme concurrentes pour l'estimation des DAP, et leurs performances respectives sont largement débattues dans la littérature internationale [1]. En France, à notre connaissance, l'intérêt de l'évaluation contingente dans le domaine de la santé a été peu étudié et celui de l'analyse conjointe encore moins.

L'objectif de notre travail consiste à explorer la capacité de l'analyse conjointe, et plus précisément celle de la Méthode des Choix Discrets (MCD), à révéler les préférences individuelles. Cette méthode consiste à présenter aux individus interrogés au moins deux options (ou « scénarios »), chaque option étant caractérisée par des niveaux d'attributs différents, et à leur demander celle qu'ils préfèrent. La MCD permet ainsi d'observer l'arbitrage que les individus effectuent, en situation de choix, entre les attributs des options qui leur sont présentées.

Nous nous proposons ici de tester la validité théorique de la méthode, en étudiant la conformité des comportements individuels à un ensemble d'hypothèses. Si ces hypothèses sont satisfaites, une fonction d'utilité peut être estimée et les taux de substitution entre attributs pris deux à deux peuvent être mesurés à partir de ses coefficients. Il est alors possible d'estimer, par exemple, les Dispositions à Payer pour faire diminuer des risques de santé.

De nombreuses recommandations ont été énoncées dans la littérature internationale, relative à la mise en œuvre de la MCD [2, 3]. On se propose donc de conduire une étude qui

respecte au mieux ces recommandations, notamment dans la phase de construction des scénarios présentés aux individus.

Le domaine d'expérimentation choisi est celui du Traitement Hormonal substitutif de la ménopause (THM). Ce traitement permet, non seulement, de réduire les troubles climatiques et le risque de fractures, mais aussi de diminuer de manière significative le risque de cancer colorectal. Malgré ses bénéfices, le THM est également associé à des risques plus élevés de cancer du sein, de maladie cardiaque et de maladie thromboembolique [4, 5, 6]. Il peut également présenter un coût monétaire à la charge de ses utilisatrices, certaines spécialités étant prises en charge par la Sécurité Sociale, d'autres pas. Il existe donc, dans cette situation, des possibilités d'arbitrage auxquelles les femmes sont confrontées. Leurs préférences et leur point de vue sur les différentes caractéristiques du traitement sont toutefois peu documentés. A notre connaissance, la méthode des choix discrets n'a jamais été mobilisée pour révéler les préférences des femmes dans ce domaine.

MATERIEL ET METHODES

La MCD s'appuie sur une enquête par questionnaire auprès d'un échantillon de la population, dont la mise en œuvre doit suivre les recommandations méthodologiques de la littérature [7]. L'étude de la validité théorique de la MCD passe par le test de certaines hypothèses de comportement individuel qui conditionnent l'existence d'une fonction d'utilité : 1) cohérence interne, 2) transitivité des préférences, 3) stabilité des préférences, 4) absence de préférences lexicographiques, 5) absence d'effet de présentation sur ces indicateurs de qualité des réponses. Le test de ces hypothèses nécessite que des questions dédiées soient élaborées au moment de la conception du questionnaire.

Construction du questionnaire

Après avoir mené une revue de la littérature relative aux différents effets du THM, nous avons produit un document de synthèse et l'avons soumis à un comité de pilotage médical composé de trois gynécologues, d'un rhumatologue et d'un cardiologue. Nous avons ensuite

réuni un groupe de discussion de sept femmes ménopausées afin de recueillir leur avis sur le THM et ses effets.

Sur la base de l'aval fourni par le comité de pilotage, et de l'analyse des commentaires du groupe de discussion, nous avons retenu 6 attributs de santé du THM, sur les critères de gravité et de fréquence, ainsi qu'un attribut de coût.

Afin de ne pas dépasser les capacités cognitives des répondantes, nous n'avons associé que deux niveaux possibles à chacun des sept attributs conservés. Les niveaux des attributs de santé correspondent 1) à leur niveau de base sans THM, et 2) à leur niveau moyen sous THM, tels qu'ils sont rapportés dans la littérature. Les niveaux de l'attribut de coût sont 1) un coût nul, qui correspond à une situation sans traitement, et 2) un coût moyen calculé à partir d'une liste des prix des THM disponibles en France actuellement, sans prise en charge par la Sécurité Sociale.

Ainsi, les attributs conservés et leurs premier et second niveaux sont respectivement :

- troubles climatiques : 70 000 vs 20 000 femmes touchées sur 100 000
- fractures liées à l'ostéoporose : 600 vs 400 femmes touchées sur 100 000 par an
- cancer colorectal : 60 vs 40 femmes touchées sur 100 000 par an
- cancer du sein : 250 vs 350 femmes touchées sur 100 000 par an
- maladie cardiaque : 150 vs 200 femmes touchées sur 100 000 par an
- maladie thromboembolique : 150 vs 350 femmes touchées sur 100 000 par an
- coût à la charge de la femme 0 vs 250 Euros par an

Les scénarios ont été construits par combinaison des attributs et niveaux d'attributs définis précédemment. A l'aide de la procédure OPTEX du logiciel SAS [8], nous avons construit un plan d'expérience fractionnaire. Ainsi, parmi les 128 (soit 2^7) combinaisons possibles, nous avons sélectionné un ensemble de huit scénarios qui respecte, d'une part, la propriété d'orthogonalité (les niveaux des attributs varient indépendamment les uns des autres) et, d'autre part, la propriété d'équilibre des niveaux (les différents niveaux de chaque attribut sont tous représentés avec la même fréquence).

Nous avons alors construit huit paires de scénarios, en associant à chacun des huit scénarios initiaux son exact opposé. Ces huit paires respectent, d'une part, la propriété d'équilibre des

utilités (aucun scénario ne domine ou n'est dominé par celui auquel il est associé) et, d'autre part, la propriété de chevauchement minimal (deux scénarios d'une même paire ne présentent aucun attribut de niveau identique). Dans la suite de l'exposé, les notations A_i et B_i désignent respectivement les premier et deuxième scénarios de la paire n° i .

Afin de tester nos hypothèses théoriques nous avons construit trois nouvelles paires de scénarios. La neuvième paire, dans laquelle A_9 domine B_9 (les niveaux de risques étant plus faible dans A_9 que dans B_9), était destinée à tester la cohérence interne. La transitivité des préférences a été testée par le biais de la dixième paire, construite telle $A_{10}=A_1$ et telle que B_1 domine B_{10} . Ainsi, si A_1 est préféré à B_1 , comme B_1 domine B_{10} , alors, par transitivité, A_{10} est également préféré à B_{10} . La stabilité des préférences des femmes a été testée par le biais de la paire n°11, répétition de la paire n°4. Nous avons détecté la présence éventuelle de préférences lexicographiques en demandant aux répondantes d'indiquer si elles avaient considéré un seul ou plusieurs attributs du THM dans leurs arbitrages entre les différents scénarios. Enfin, afin de tester l'effet de présentation de l'information sur les préférences des femmes, deux versions du questionnaire ont été construites. Dans la première, tous les attributs du traitement étaient exprimés pour 100 000 femmes. Dans la seconde, les ordres de grandeur variaient d'un attribut à l'autre.

Pour analyser les facteurs explicatifs du point de vue des femmes et de leurs préférences, nous avons également recueilli leurs caractéristiques sociodémographiques et médicales.

Le questionnaire a été pré-testé auprès de 20 femmes, pré-ménopausées ou ménopausées, sous THM ou non, pour moitié en face à face, et pour moitié lors d'un entretien téléphonique. Les femmes dont les réponses ont été recueillies par téléphone avaient reçu le questionnaire par voie postale au préalable.

Passation des questionnaires

Les répondantes ont été recrutées par téléphone à partir de listes des abonnés téléphoniques de Lyon et sa région, sur les critères suivants: être dans la capacité de lire le français, âgée de 45 à 65 ans, pré-ménopausée ou ménopausée, sous THM ou non. Le questionnaire a été envoyé, par voie postale, aux femmes qui avaient donné leur accord au

téléphone. Les répondantes pouvaient ensuite renvoyer le questionnaire une fois rempli à l'aide d'une enveloppe jointe, ou attendre qu'un enquêteur les rappelle pour collecter leurs réponses par téléphone. Dans les deux cas, elles devaient avoir rempli leur questionnaire au préalable. Le taux de non réponse a été limité en effectuant jusqu'à deux relances téléphoniques.

Analyse des réponses

Les réponses ont été analysées à l'aide de statistiques descriptives.

Plus précisément, la cohérence interne a été testée en calculant la proportion des femmes qui avaient préféré le scénario dominé au scénario dominant dans la paire construite spécifiquement pour ce test. L'hypothèse de transitivité n'a pu être testée que parmi la sous-population des femmes qui avaient préféré le scénario A dans la première paire de scénarios et/ou qui avaient préféré le scénario B dans la dixième paire. Nous avons alors calculé la proportion des femmes qui avaient fourni une réponse identique à ces deux paires. L'hypothèse de stabilité des préférences a été testée en calculant la proportion des femmes qui exprimaient une préférence identique aux quatrième et onzième paires de scénarios.

La recherche de préférences lexicographiques a d'abord consisté à identifier les femmes qui avaient indiqué n'avoir pris en compte qu'un unique attribut dans les onze choix successifs. Nous nous sommes ensuite assuré que, dans le scénario qu'elles avaient choisi à chacune des onze paires, l'attribut qu'elles avaient désigné comme le seul pris en compte présentait bien son niveau le plus favorable.

L'effet de présentation des risques sur la qualité des réponses a été analysé par le biais de tests du chi-deux, entre la variable de présentation et les différents indicateurs de qualité (cohérence interne, transitivité, stabilité, absence de préférences lexicographiques).

La prochaine étape de l'analyse des réponses consistera à analyser les préférences des femmes, par le biais de modèles de régression logistiques. Il s'agira d'estimer la variation d'utilité retirée par les femmes lorsqu'elles choisissent le bien décrit par le scénario A plutôt que le bien décrit par le scénario B.

Rappelons que chaque femme a effectué 8 choix successifs proposés à partir de l'ensemble des huit paires de scénarios originales qui respectent les critères d'orthogonalité, d'équilibre des niveaux, de chevauchement minimal, et d'équilibre des utilités. Par conséquent, afin de prendre en compte les éventuelles corrélations entre les 8 observations issues d'une même répondante, nous aurons recours à une modélisation à deux niveaux à l'aide du logiciel HLM6 [9]. La variable dépendante sera une variable dichotomique qui reflète la préférence exprimée par chaque femme pour chacune des huit paires. Les variables explicatives de niveau 1 seront les différences de niveaux des attributs entre les scénarios A et B. Des effets aléatoires seront introduits dans le modèle afin de déterminer si les effets des attributs du THM sont constants ou varient entre les femmes. Les variables explicatives de niveau 2 seront les variables sociodémographiques et médicales des femmes énoncées plus haut, dont nous pouvons supposer qu'elles ont une influence sur l'importance relative qu'elles accordent aux différents attributs des THM lorsqu'elles prennent la décision de suivre ou non un tel traitement.

RESULTATS

Sur la base des listes des abonnés téléphoniques, 4534 femmes habitant Lyon et sa région ont été sollicitées pour participer à l'étude. Parmi elles, 669 femmes satisfaisaient les critères d'inclusion et ont accepté de participer à l'étude. Finalement, 462 (soit 69%) ont dûment rempli leur questionnaire.

Description de la population

L'âge moyen des femmes interrogées était de 56 ans. Plus de la moitié d'entre elles vivaient en couple (52%) et exerçaient une activité rémunérée (54%). Elles étaient une grande majorité à bénéficier d'une assurance-maladie complémentaire (95%). Plus de la moitié des femmes n'avaient pas suivi d'études supérieures (56%), et avaient un niveau de revenu annuel avant impôts inférieur à 45 000€ (70%). Enfin, 64% participaient à des activités extraprofessionnelles.

Trois cent onze femmes (67%) étaient ménopausées. Parmi elles, 198 suivaient ou avaient déjà suivi un THM. Quatre-vingt-trois pourcent des femmes déclaraient bénéficier d'un

suivi gynécologique régulier. En ce qui concerne leurs antécédents relatifs aux pathologies pouvant être liées au THM, 34% présentaient des antécédents de maladie cardiaque, 30% présentaient des antécédents de cancer du sein, 28% présentaient des antécédents d'ostéoporose, 23% présentaient des antécédents de maladie thromboembolique et 15% présentaient des antécédents de cancer colorectal.

Test des hypothèses de comportement

Parmi les 462 répondantes, 441 (soit 95%, Intervalle de Confiance à 95% = [0.90; 0.99]) ont indiqué qu'elles préféraient le scénario dominant au scénario dominé dans la paire construite pour ce test. Seules 5 % des femmes interrogées ont donc répondu de manière incohérente. Nous avons testé l'hypothèse de transitivité des préférences parmi les 219 patientes qui avaient déclaré préférer le scénario A à la première paire, et/ou qui avaient déclaré préférer le scénario B à la dixième paire. Sur ces 219 femmes, 213 (97% IC₉₅ = [0.90 ; 1.00]) ont présenté des préférences transitives, soit que, ayant choisi A à la première paire, elles aient aussi choisi A à la dixième paire, ou que, ayant choisi B à la dixième paire, elles aient aussi choisi B à la première paire.

Quatre cent sept femmes (soit 88%, IC₉₅ = [0.83 ; 0.92]) ont exprimé des préférences stables, en fournissant la même réponse aux quatrième et onzième paires de scénarios. Le coefficient Kappa, qui compare la concordance observée entre les réponses à ces deux paires et la concordance aléatoire s'élève à 0.62, ce qui, selon Landis & Koch [9] indique une bonne concordance.

Soixante-dix-huit femmes (17%, IC₉₅ = [0.12 ; 0.21]) ont déclaré n'avoir fondé leurs préférences que sur un attribut unique. Les attributs les plus fréquemment cités comme seul critère de comparaison des scénarios étaient: 1) les troubles climatiques (cités par 33% des femmes), 2) le coût du traitement (24%), et 3) le risque de cancer du sein (22%). Venaient ensuite le risque cardiaque (7%), les fractures liées à l'ostéoporose (5%), le risque de cancer colorectal (4%) et le risque thromboembolique (4%). L'exploitation des préférences exprimées par ces 78 femmes à chacune des onze paires de scénarios a toutefois montré que seules 57 d'entre elles avaient effectivement toujours choisi le scénario dans lequel l'attribut cité comme

unique critère de comparaison présentait son niveau le plus favorable. Le taux de préférences réellement lexicographiques s'élève donc à 12%.

Les p-value des tests du chi-deux entre la variable de présentation des risques et les résultats des tests des hypothèses de comportement (cohérence interne, transitivité et stabilité des préférences, préférences lexicographiques) sont toutes supérieures à 5%. Ainsi, il semblerait que la présentation des risques n'ait pas d'effet sur la qualité des réponses.

DISCUSSION /CONCLUSION

Les répondantes semblent avoir eu une bonne compréhension du questionnaire et de l'arbitrage qui leur était demandé. En effet, face à la paire de scénarios destinée à tester la cohérence interne, seules 5% des femmes ont indiqué qu'elles préféraient le scénario dominé au scénario dominant. Ce résultat nous semble très satisfaisant car, à titre de comparaison, dans leur étude relative aux préférences des femmes pour la prise en charge d'une fausse couche, Ryan & Hugues [11] avaient observé que 3 à 38% des femmes avaient répondu de manière cohérente.

Le taux de répondantes qui expriment des préférences stables est également satisfaisant puisqu'il s'élève à 88%. Dans leur étude relative à la prise en charge des blessures au genou, Bryan *et.al*/ [12] avaient testé la stabilité des préférences des répondants en répétant quatre des huit paires de choix originales. Le taux de répondants stables sur les 4 questions s'élevait à 57% seulement. Enfin, seulement douze pourcent des répondantes ont exprimé des préférences lexicographiques alors que 70% des femmes interrogées par Ratcliffe & Longworth [13] avaient exprimé une telle préférence pour l'un des attributs de soins postnataux sur lesquels portait l'enquête.

Ainsi, les comportements individuels sont le plus souvent conformes aux hypothèses : bonne cohérence interne, faible taux de préférences instables, de préférences intransitives et de préférences lexicographiques, absence d'effet de présentation sur la qualité des réponses. Il sera donc possible de poursuivre l'analyse en étudiant les arbitrages des femmes entre les

différents attributs du THM. Dans un premier temps, l'estimation des coefficients des attributs dans la fonction d'utilité permettra de comparer leurs signes aux sens attendus, notamment grâce à la littérature empirique sur le THM. La présence d'un attribut de coût permettra d'étudier les arbitrages entre coût et risque ou bénéfice de santé. On pourra également rechercher la présence d'un effet de présentation des risques (*framing effects*), non plus sur la qualité des réponses fournies par les femmes, mais sur leurs préférences. Enfin, on pourra étudier la sensibilité de la préférence des femmes à leurs caractéristiques personnelles, sociodémographiques et médicales.

REFERENCES

1. Cookson R. Willingness to pay methods in health care: a sceptical view. *Health Econ* 2003;12:891-894
2. Zwerina K, Huber J, Kuhfeld W. *A general method for constructing efficient choice designs. Dans: Marketing Research Methods in SAS*. SAS 9.1 Edition. 2005:49 pages.
3. Louviere JJ, Hensher DA, Swait JD. *Stated Choice Methods: Analysis and Application*, Cambridge, U.K: Cambridge University Press. 2000:418 pages
4. Writing group for the women's health initiative investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. *JAMA* 2002;288(3):321-368.
5. Million women study collaborators. Breast cancer and hormone-replacement therapy in the million women study. *Lancet* 2003;362:419-427.
6. Fournier A, Berrino F, Riboli E, Avenel V, Clavel-Chapelon F. Breast cancer risk in relation to different types of hormone replacement therapy in the E3N-EPIC cohort. *Int J Cancer* 2005;114:448-454.
7. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ* 2000;320:1530-3
8. Sas Institute Inc *SAS/STAT User's Guide Version 8*, 2nd edition, Cary, NC: Stats Publishing Inc. 1999: 1464 pages

9. Raudenbush S, Bryk A, Cheong YF, Congdon R. *HLM 6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: Scientific Software International. 2004:310pages.
10. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 1977;33:159-174.
11. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. *Health Econ* 1997;6(3):261-73.
12. Bryan S, Gold L, Sheldon R, Buxton M. Preference measurement using conjoint methods: an empirical investigation of reliability. *Health Econ* 2000;9(5):385-95
13. Ratcliffe J, Longworth L. Investigating the structural reliability of a discrete choice experiment within health technology assessment. *Int Journal Technol Assess Health Care* 2002;18(1):139-44.

Remerciements : Florence Nguyen est allocataire de la Fondation de France